AD-A097 619 PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS F/G 12/1
EMPIRICAL BAYES RULES FOR SELECTING GOOD POPULATIONS.(U)
MAR 81 S S GUPTA, P HSIAO N00014-75-C-0455
UNCLASSIFIED MIMEOGRAPH-SER-81-5 NL

1 OF 1
AD
A097619

END
DATE
FILMED
5 81
DTIC

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

Empirical Bayes Rules for
Selecting Good Populations*

by

Shanti S. Gupta
Purdue University

and

Ping Hsiao
Wayne State University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #81-5

March, 1981

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Empirical Bayes Rules for
Selecting Good Populations*

by

Shanti S. Gupta
Purdue University

and

Ping Hsiao
Wayne State University

1. Introduction

We assume that G is an unknown prior distribution on Θ, and denote the minimum Bayes risk in a decision problem by r(G). Robbins, in his pioneering papers [1955], [1964], proposed sequences of decision rules, based on data from n independent repetitions of the same decision problem, whose (n+1)st stage Bayes risk converges to r(G) as $n \to \infty$. Such sequences of rules are called empirical Bayes rules. Empirical Bayes rules have been derived for multiple decision problems by Deely [1965], Van Ryzin [1970], Huang [1975], Van Ryzin and Susarla [1977], and Singh [1977]. However, the forms of densities of the populations that these authors considered are either $c(\theta)h(x)e^{\theta x}$, for continuous case or $c(\theta)h(x)\theta^x$, for discrete case, and the loss function are either squared error or merely $\max_{1\le j\le k} \theta_j-\theta_i$ type. Fox [1978] discussed some estimation problem under squared error loss, in which empirical Bayes rules were derived for uniform distributions for the first time. Barr and Rizvi [1966], and McDonald [1974] also considered selection problems related to uniform distribution by subset selection approach. The problem considered in this paper is related to uniform distributions and can

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

be illustrated by the following example. Suppose that there are k drugs for a certain disease, and the effect of drug i follows an unknown distribution G_i, $1 \le i \le k$. The effectiveness of drug i is tested on n patients (For different drugs, different groups of patients are used. If the same patient has to be used for more than one test, let there be a wash-out period between tests, so the effects of different drugs are independent.). Let θ_{ij} be a measurement of the effectiveness of drug i on patient j. Drug i cures the disease of patient j if $\theta_{ij} \ge \theta_0$ and hence is entitled as a good drug, otherwise it is a bad drug. θ_0 is called the control parameter. In general, θ_{ij} is unknown and will diminish gradually as time passed by, so a diagnosis will yield a result Y_{ij} which we assume to be uniformly distributed over $(0, \theta_{ij})$. Our purpose is to decide on the quality (good or bad) of the k drugs on the next consulting patient based on Y_{ij} $(1 \le i \le k,\ 1 \le j \le n)$ and X_i $(1 \le i \le k)$, where X_i is the diagnostic result of drug i on the present patient. In Section 2, a general formulation is given and empirical Bayes rules are derived for selecting populations better than a known control when the populations are uniformly distributed. In Section 3, the same problem is considered except that the control parameter is unknown. In Section 4, empirical Bayes rules are found for truncation parameters (that is the densities are of the form $p_i(x)c_i(\theta_i)I_{(0,\theta_i)}(x)$). Rate of convergence is also discussed. Monte Carlo studies are carried out for the priors $G(\theta) = \frac{\theta^2}{c^2} I_{(0,c)}(\theta)$. The smallest sample size N is determined to guarantee that the relative error is less than ε.

2. Known control parameter

Assume that $\pi_1, \pi_2, \ldots, \pi_k$ are k populations and $\pi_i \sim U(0,\theta_i)$, where θ_i is unknown for $1 \le i \le k$. Let θ_0 be a known control parameter, we define π_i

to be a good population if $\theta_i \geq \theta_0$ and to be a bad population if $\theta_i < \theta_0$. Let $\Omega = \{\theta = (\theta_1,\ldots,\theta_k) \mid \theta_i > 0$ for all $1 \leq i \leq k\}$. For any $\theta \in \Omega$, let $A(\theta) = \{i \mid \theta_i \geq \theta_0\}$ and $B(\theta) = \{i \mid \theta_i < \theta_0\}$, then $A(\theta)$ $(B(\theta))$ is the set of indices of good (bad) populations. Our goal is to select all the good populations and reject the bad ones. We formulate the problem in the empirical Bayes framework as follows:

(1) Let $\mathcal{A} = \{S \mid S \subset \{1,2,\ldots,k\}\}$ be the action space. When we take action S, we say π_i is good if $i \in S$ and π_i is bad if $i \notin S$.

(2) $L(\theta,S) = L_1 \sum_{i \in A(\theta)\setminus S} (\theta_i-\theta_0) + L_2 \sum_{i \in B(\theta)\cap S} (\theta_0-\theta_i)$ is the loss function. (2.1)

(3) Let $dG(\theta) = \prod_{i=1}^{k} dG_i(\theta_i)$ be an unknown prior distribution on Ω, where G_i has a continuous pdf g_i.

(4) Let $(\theta_{i1},Y_{i1}),\ldots,(\theta_{in},Y_{in})$ be pairs of random variables from π_i and $Y_{ij} \mid \theta_{ij} \sim U(0,\theta_{ij})$ for $1 \leq i \leq k$, $1 \leq j \leq n$. Let $Y_j = (Y_{1j},\ldots,Y_{kj})$, then Y_j denotes the previous j-th observations from $\pi_1,\ldots,\pi_k$.

(5) Let $X = (X_1,\ldots,X_k)$ be the present observation and $f(x|\theta) = \prod_{i=1}^{k} \frac{1}{\theta_i} I_{(0,\theta_i)}(x_i)$. Since we are interested in Bayes rules, we can restrict our attention to the non-randomized rules.

(6) Let $D = \{\delta : \mathcal{X} \to \mathcal{A}$ is measurable$\}$, then $r(G) = \inf_{\delta \in D} r(G,\delta)$ is the minimum Bayes risk.

The decision rules $\{\delta_n(x; Y_1,\ldots,Y_n)\}_{n=1}^{\infty}$ is said to be asymptotically optimal (a.o.) or empirical Bayes (e.B.) relative to G if $r_n(G,\delta_n) = \int_{\mathcal{X}} E \int_{\Omega} L(\theta,\delta_n(x,Y_1,\ldots,Y_n)) f(x|\theta)\,dG(\theta)\,dx \to r(G)$ as $n \to \infty$. For simplicity,

$\delta_n(x, Y_1, \ldots, Y_n)$ will be denoted by $\delta_n(\underset{\sim}{x})$.

Let $m_i(x)$ be the marginal pdf of X_i and $M_i(x)$ be the marginal distribution of X_i. Then we have

$$m_i(x) = \int_x^\infty \frac{1}{\theta}\, dG_i(\theta) \quad \text{for all } x > 0, \tag{2.2}$$

and

$$M_i(x) = \int_0^x \int_t^\infty \frac{1}{\theta}\, dG_i(\theta)dt = \int_x^\infty \int_0^x \frac{1}{\theta}\, dtdG_i(\theta) + \int_0^x \int_0^\theta \frac{1}{\theta}\, dtdG_i(\theta)$$

$$= xm_i(x) + G_i(x).$$

Hence, $G_i(x) = M_i(x) - xm_i(x)$. $\qquad$ (2.3)

Now, the loss function defined in (2.1) can be expressed as

$$L(\theta, S) = \sum_{i \in S} [L_2(\theta_0 - \theta_i) I_{(0,\theta_0]}(\theta_i) - L_1(\theta_i - \theta_0) I_{(\theta_0,\infty)}(\theta_i)]$$

$$+ \sum_{i=1}^{k} L_1(\theta_i - \theta_0) I_{(\theta_0,\infty)}(\theta_i). \tag{2.4}$$

The second sum in (2.4) does not depend on the action S. To find the Bayes rule we can omit it, and only consider the first sum as our loss from now on. Then,

$$r(G, \delta) = \int_{\mathcal{X}} \sum_{i \in \delta(\underset{\sim}{x})} [\int_{\theta_i \le \theta_0} L_2(\theta_0 - \theta_i) f(\underset{\sim}{x} | \underset{\sim}{\theta}) dG(\underset{\sim}{\theta})$$

$$- \int_{\theta_i > \theta_0} L_1(\theta_i - \theta_0) f(\underset{\sim}{x} | \underset{\sim}{\theta}) dG(\underset{\sim}{\theta})] d\underset{\sim}{x}.$$

So, if $\delta_B(\underset{\sim}{x}) = S^*$ is the Bayes rule, one finds $i \in S^*$ if

$$\int_{(0,\theta_0] \cap (x_i,\infty)} L_2(\theta_0 - \theta_i) \frac{1}{\theta_i}\, dG_i(\theta_i)$$

$$< \int_{\theta_0 \vee x_i}^{\infty} L_1(\theta_i - \theta_0) \frac{1}{\theta_i}\, dG_i(\theta_i). \quad \text{Hence,}$$

$S^* = \{i \mid x_i \geq \theta_0\} \cup \{i \mid x_i < \theta_0 \text{ and } H_i(x_i) \leq c_i(\theta_0)\}$ where

$$H_i(x_i) = L_2\theta_0 \int_{x_i}^{\theta_0} \frac{1}{\theta_i}\, dG_i(\theta_i) + L_2G_i(x_i) \quad \text{and}$$

$$c_i(\theta_0) = L_2G_i(\theta_0) + L_1(1-G_i(\theta_0)) - L_1\theta_0 \int_{\theta_0}^{\infty} \frac{1}{\theta_i}\, dG_i(\theta_i).$$

Since $H_i(x_i)$ is decreasing in x_i for $x_i < \theta_0$ and $H(\theta_0) \leq c_i(\theta_0)$, so $S^* = \{i \mid x_i \geq \theta_0 - b_i\}$ where $b_i \geq 0$ satisfies $H(\theta_0 - b_i) = c_i(\theta_0)$. This shows for any G, Gupta type rules are Bayes rules (see Gupta [1958,1963,1965]). Now, since G is unknown, the Bayes rules are not obtainable. We wish to find a sequence of rules $\{\delta_n(x)\}_{n=1}^{\infty}$ to be a.o. Let

$$\Delta_{G_i}(x_i) = H_i(x_i) - c_i(\theta_0)$$

and

$$S_B(x) = \{i \mid x_i < \theta_0,\ \Delta_{G_i}(x_i) \leq 0\}.$$

Also, for any i $(1 \leq i \leq k)$, let $\Delta_{i,n}(x_i) = \Delta_i(x_i, Y_{i1}, \ldots, Y_{in})$ for all $n = 1,2,\ldots,$ be a sequence of real-valued measurable functions, we define

$$S_n(x) = \{i \mid x_i < \theta_0 \text{ and } \Delta_{i,n}(x_i) \leq 0\} \tag{2.5}$$

and

$$\delta_n(x) = \{i \mid x_i \geq \theta_0\} \cup S_n(x). \tag{2.6}$$

One can show that

Theorem 2.1. If $\int_0^{\infty} \theta\, dG_i(\theta) < \infty$ $\forall\ i = 1,2,\ldots,k$, and $\Delta_{i,n}(x_i) \to \Delta_{G_i}(x_i)$ in (p) for almost all $x_i < \theta_0$. Then $\{\delta_n(x)\}_{n=1}^{\infty}$ defined by (2.6) is e.B.

Proof: For all $S \subset \Omega$, let

$$\mathcal{X}_S = \{x \mid x_i \geq \theta_0 \text{ if } i \in S \text{ and } x_i < \theta_0 \text{ if } i \notin S\}.$$

Now, for any $x \in \mathcal{X}_S$, $\delta_B(x) = S \cup S_B(x)$, then

$$\int L(\theta, \delta_B(x)) f(x|\theta) dG(\theta)$$

$$= \sum_{i \in \delta_B(x)} \left[\int_{\theta_i < \theta_0} L_2(\theta_0 - \theta_i) f(x|\theta) dG(\theta) - \int_{\theta_i > \theta_0} L_1(\theta_i - \theta_0) f(x|\theta) dG(\theta)\right]$$

$$= \sum_{i \in S} - Q(x) + \sum_{i \in S_B(x)} \Delta_{G_i}(x_i) \prod_{j \neq i} m_j(x_j)$$

where $Q(x) = \int_{\theta_i > \theta_0} L_1(\theta_i - \theta_0) f(x|\theta) dG(\theta).$

Similarly, for $x \in \mathcal{X}_S$, we have

$$\int_{\Theta} L(\theta, \delta_n(x)) f(x|\theta) dG(\theta)$$

$$= \sum_{i \in S} - Q(x) + \sum_{i \in S_n(x)} \Delta_{G_i}(x_i) \prod_{j \neq i} m_j(x_j).$$

Hence, if $\Delta_{i,n}(x_i) \to \Delta_{G_i}(x_i)$ in (p), then

$$0 \leq \int_{\Theta} [L(\theta, \delta_n(x)) - L(\theta, \delta_B(x))] f(x|\theta) dG(\theta)$$

$$\leq \sum_{i \in S_n(x)} |\Delta_{G_i}(x_i) - \Delta_{i,n}(x_i)| \prod_{j \neq i} m_j(x_j)$$

$$+ \left(\sum_{i \in S_n(x)} - \sum_{i \in S_B(x)}\right) \Delta_{i,n}(x_i) \prod_{j \neq i} m_j(x_j) \tag{2.7}$$

$$+ \sum_{i \in S_B(x)} |\Delta_{i,n}(x_i) - \Delta_{G_i}(x_i)| \prod_{j \neq i} m_j(x_j)$$

$$\leq 2\varepsilon \sum_{i=1}^{k} \prod_{j \neq i} m_j(x_j)$$

with probability near 1 for $n > N$. Note that (2.7) is non-positive by the definition of $S_n(x)$. Thus, we have proved

$$\int_{\Theta} L(\theta,\delta_n(\underset{\sim}{x}))f(\underset{\sim}{x}|\underset{\sim}{\theta})dG(\underset{\sim}{\theta}) \to \int_{\Theta} L(\underset{\sim}{\theta},\delta_B(\underset{\sim}{x}))f(\underset{\sim}{x}|\underset{\sim}{\theta})dG(\underset{\sim}{\theta})$$

in (p) for all most all x. By Corollary 1 of Robbins [1964], $\{\delta_n(\underset{\sim}{x})\}_{n=1}^{\infty}$ is e.B. This completes the proof.

In view of (2.2) and (2.3), we have

$$\Delta_{G_i}(x_i) = L_2 m_i(x_i)(\theta_0-x_i) + L_2[M_i(x_i)-M_i(\theta_0)] + L_1[M_i(\theta_0)-1].$$

Hence, if we define

$$\Delta^*_{i,n}(x_i) = L_2 m_{i,n}(x_i)(\theta_0-x_i) + L_2[M_{i,n}(x_i)-M_{i,n}(\theta_0)] + L_1[M_{i,n}(\theta_0)-1] \tag{2.8}$$

where $\quad M_{i,n}(x) = \frac{1}{n}\sum_{j=1}^{n} I_{(-\infty,x]}(Y_{ij})$

and $\quad m_{i,n}(x) = \frac{1}{h}[M_{i,n}(x+h) - M_{i,n}(x)], \qquad (2.9)$

then $\Delta^*_{i,n}(x_i) \to \Delta_{G_i}(x_i)$ in (p) a.e. in $\underset{\sim}{x}$, if $h = h(n) \to 0$ and $nh \to \infty$ as $n \to \infty$. So, by Theorem 1, $\delta^*_n(\underset{\sim}{x}) = \{i|x_i \geq \theta_0\} \cup \{i|x_i < \theta_0, \Delta^*_{i,n}(x_i) \leq 0\}$ is e.B.

Remark: In (2.8), $M_{i,n}(x)$ and $m_{i,n}(x)$ can be defined as any functions such that $M_{i,n}(x) \to M_i(x)$ in (p) and $m_{i,n}(x) \to m_i(x)$ in (p) for almost all x. For example, let $m^0_{i,n}(x) = \frac{1}{nh}\sum_{j=1}^{n} w(\frac{x-Y_{ij}}{h})$ where $w(\cdot) \geq 0$ satisfies

(i) $\sup_{-\infty<x<\infty} w(x) < K$ for some constant K,

(ii) $\int_{-\infty}^{\infty} w(x)dx = 1$

(iii) $\lim_{x\to\infty} xw(x) = 0$

and $h = h(n)$ satisfies $h \to 0$, $nh \to \infty$ as $n \to \infty$ then $m^0_{i,n}(x)$ is a consistent estimator of $m_i(x)$ (see Parzen [1962]).

3. θ_0 unknown

Let π_0 be a control population and $\pi_0 \sim U(0,\theta_0)$ with θ_0 unknown. Let $Y_{01},\ldots,Y_{0n}$ be the past data collected from π_0. Based on this further information, we will search for empirical Bayes rules for selecting populations better than control. Note that now $\underset{\sim}{\theta} = (\theta_0,\theta_1,\ldots,\theta_k)$, $\underset{\sim}{x} = (x_0,x_1,\ldots,x_k)$ and $G(\cdot) = \prod_{i=0}^{k} G_i(\theta_i)$. Under the loss function in (2.4), the Bayes rule δ_B is: $i \in \delta_B(x)$ if

$$L_2 \int_{x_0}^{\infty} \frac{1}{\theta_0} \int_{(0,\theta_0]\cap(x_i,\infty)} \frac{1}{\theta_i} (\theta_0-\theta_i)dG_i(\theta_i)dG_0(\theta_0)$$

$$\le L_1 \int_{x_0}^{\infty} \frac{1}{\theta_0} \int_{(\theta_0,\infty)\cap(x_i,\infty)} \frac{1}{\theta_i} (\theta_i-\theta_0)dG_i(\theta_i)dG_0(\theta_0).$$

Hence, $i \in \delta_B(x)$ if

(i) $x_i \ge x_0$ and $\Delta^1_{G_0,G_i}(x_0,x_i) \le 0$, where

$$\Delta^1_{G_0,G_i}(x_0,x_i) = (L_1-L_2)[\int_{x_i}^{\infty} m_i(\theta_0)dG_0(\theta_0) + \int_{x_i}^{\infty} m_0(\theta_i)dG_i(\theta_i)]$$

$$- L_1[1-G_i(x_i)]m_0(x_0)+m_i(x_i)[L_2+(L_1-L_2)G_0(x_i)-L_1G_0(x_0)] \qquad (3.1)$$

or

(ii) $x_i < x_0$ and $\Delta^2_{G_0,G_i}(x_0,x_i) \le 0$, where

$$\Delta^2_{G_0,G_i}(x_0,x_i) = (L_1-L_2)[\int_{x_0}^{\infty} m_i(\theta_0)dG_0(\theta_0) + \int_{x_0}^{\infty} m_0(\theta_i)dG_i(\theta_i)]$$

$$- m_0(x_0)[L_1+(L_2-L_1)G_i(x_0)-L_2G_i(x_i)] + L_2m_i(x_i)(1-G_0(x_0)). \qquad (3.2)$$

When $L_1 = L_2 = L$, the Bayes rule is greatly simplified. We find $i \in \delta_B(x)$ if

$$\Delta_{G_0,G_i}(x_0,x_i) = m_0(x_0)[1-G_i(x_i)]-m_i(x_i)[1-G_0(x_0)] \geq 0.$$

Let $\delta_n(x) = \{i \mid \Delta_{i,n}(x_i,x_0) \geq 0\}$ where

$$\Delta_{i,n}(x_i,x_0) = m_{0,n}(x_0)[1-G_{i,n}(x_i)] - m_{i,n}(x_i)[1-G_{0,n}(x_0)],$$

$m_{i,n}(x_i)$ is defined in (2.9), and $G_{i,n}(x_i) = M_{i,n}(x_i) - x_i m_{i,n}(x_i)$. Then, $\{\delta_n(x)\}_{n=1}^{\infty}$ is e.B. by Theorem 3.2. When $L_1 \neq L_2$, one needs to find consistent estimators of $\int_a^\infty m_i(\theta_0)dG_0(\theta_0)$ and $\int_a^\infty m_0(\theta_i)dG_i(\theta_i)$.

Theorem 3.1. Let $M_{i,n}(x)$ and $m_{i,n}(x)$ be defined by (2.9) with $h = h(n)$ satisfying $h \to 0$, $nh^2 \to \infty$ as $n \to \infty$. If $\int_0^\infty \theta \, dG_i(\theta) < \infty$ for all $i = 0,1,\ldots,k$, then $\int_a^\infty m_{i,n}(x)dm_{0,n}(x) \to \int_a^\infty m_i(x)dG_0(x)$ in (p) for any $a > 0$.

Proof: See Appendix A.

Theorem 3.2. Assume that $\int_0^\infty \theta \, dG_i(\theta) < \infty$ for all $0 \leq i \leq k$. If for all $1 \leq i \leq k$, $\Delta^1_{i,n}(x_0,x_i) \to \Delta^1_{G_i,G_0}(x_0,x_i)$ in (p) for $x_i \leq x_0$, and $\Delta^2_{i,n}(x_0,x_i) \to \Delta^2_{G_i,G_0}(x_0,x_i)$ in (p) for $x_i > x_0$. Then

$$\begin{aligned} \delta^*_n(x) &= \delta^1_n(x) \cup \delta^2_n(x) \\ &= \{i \mid x_i \leq x_0 \text{ and } \Delta^1_{i,n}(x_0,x_i) \geq 0\} \cup \\ &\quad \{i \mid x_i > x_0 \text{ and } \Delta^2_{i,n}(x_0,x_i) \geq 0\} \end{aligned} \tag{3.3}$$

defines an empirical Bayes rule.

Proof: $\int_{\Omega} L(\theta,\delta_B(\underset{\sim}{x}))f(\underset{\sim}{x}|\theta)dG(\theta)$

$$= \sum_{i\in S_1^\star(\underset{\sim}{x})} \Delta^1_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j) + \sum_{i\in S_2^\star(\underset{\sim}{x})} \Delta^2_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j)$$

where $S_1^\star(\underset{\sim}{x}) = \{i|x_i > x_0 \text{ and } \Delta^1_{G_i,G_0}(x_0,x_i) \leq 0\}$

$S_2^\star(\underset{\sim}{x}) = \{i|x_i < x_0 \text{ and } \Delta^2_{G_i,G_0}(x_0,x_i) \leq 0\}$,

and $\int_{\Omega} L(\theta,\delta_n^\star(\underset{\sim}{x}))f(\underset{\sim}{x}|\theta)dG(\theta)$

$$= \sum_{i\in S_n^1(\underset{\sim}{x})} \Delta^1_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j) + \sum_{i\in S_n^2(\underset{\sim}{x})} \Delta^2_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j).$$

Now, following the same method as in the proof of Theorem 2.1, we can show

$$\sum_{i\in S_n^\ell(\underset{\sim}{x})} \Delta^\ell_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j) \to \sum_{i\in S_\ell^\star(\underset{\sim}{x})} \Delta^\ell_{G_i,G_0}(x_0,x_i)\prod_{j\neq i} m_j(x_j)$$

in (p) for $\ell = 1,2$. Hence $\{\delta_n^\star(\underset{\sim}{x})\}_{n=1}^\infty$ is e.B. This completes the proof.

Now, let

$$\Delta^1_{i,n}(x_0,x_i) = (L_2-L_1)\{\int_{x_i}^\infty xm_{i,n}(x)dm_{0,n}(x) + \int_{x_i}^\infty xm_{0,n}(x)dm_{i,n}(x)\}$$
$$- L_1[1-G_{i,n}(x_i)]m_{0,n}(x_0)+m_{i,n}(x_i)[L_2+(L_1-L_2)$$
$$G_{0,n}(x_i)-L_1G_{0,n}(x_0)], \qquad (3.4)$$

and

$$\Delta^2_{i,n}(x_0,x_i) = (L_2-L_1)\{\int_{x_0}^\infty xm_{i,n}(x)dm_{0,n}(x) + \int_{x_0}^\infty xm_{0,n}(x)dm_{i,n}(x)\}$$
$$+ L_2[1-G_{0,n}(x_0)]m_{i,n}(x_i)-m_{0,n}(x_0)[L_1+(L_2-L_1)G_{i,n}(x_0)$$
$$- L_2G_{i,n}(x_i)],$$

where $G_{i,n}(x) = M_{i,n}(x) - xm_{i,n}(x)$. (3.5)

Then, by Theorem 3.1 and Theorem 3.2 (3.3), (3.4), and (3.5) define an empirical Bayes rule.

4. Generalization and Simulation

Let $p_i(x)$ be a positive continuously differentiable function which is defined over $(0,\infty)$ for $1 \leq i \leq k$. Let $c_i(\theta)^{-1} = \int_0^\theta p_i(x)dx$ for $\theta > 0$, then $f_i(x|\theta) = p_i(x)c_i(\theta)I_{(0,\theta)}(x)$ is a density function and θ is a truncation parameter. In this section, we assume that $\pi_i \sim f_i(x|\theta_i)$ for $1 \leq i \leq k$. Under the formulation of Section 2, we wish to find empirical Bayes rules for these more general density functions. For simplicity, we assume that $L_1 = L_2 = L$ and that θ_0 is known. Also we assume $G_i(\theta)$ has a continuous density $g_i(\theta)$ with a bounded support $[0,\alpha_i]$ with a known α_i for all $1 \leq i \leq k$. We find

$$m_i(x) = \int_0^{\alpha_i} f_i(x|\theta)dG_i(\theta) = p_i(x)\int_x^{\alpha_i} c_i(\theta)dG_i(\theta).$$

If we follow the same discussion as in Section 2, we can show that the Bayes rule δ_B is $i \in \delta_B(\underset{\sim}{x})$ iff

(i) $x_i \geq \theta_0$, or

(ii) $x_i < \theta_0$ and $\theta_0 \int_x^{\alpha_i} c_i(x)dG_i(x) \leq \int_{x_i}^{\alpha_i} xc_i(x)dG_i(x)$.

Hence, $\delta_B(\underset{\sim}{x}) = \{i|x_i \geq \theta_0 - d_i\}$ where $d_i \geq 0$ satisfies $\int_{d_i}^{\alpha_i} (\theta_0 - x)c_i(x)dG_i(x) = 0$. Let $d_{i,n} = d_{i,n}(Y_{i1},\ldots,Y_{in})$ be a consistent estimation of d_i, then $\delta_n^0(\underset{\sim}{x}) = \{i|x_i \geq \theta_0 - d_{i,n}\}$ is e.B. and they are (weak) admissible in the sense that $\delta_n^0(\cdot,\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n)$ is an admissible rule for the non-empirical problem for all $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n$ and n (see Houwelingen (1976). Meeden (1972)). However,

to find such a sequence $\{d_{i,n}\}_{n=1}^{\infty}$ is very difficult. In view of Theorem 2.1, a more practical way to find empirical Bayes rules is to estimate

$$\int_{x_i}^{\alpha_i} xc_i(x)dG_i(x).$$

Theorem 4.1. Let $p_i(x)$ and $G_i(x)$ be defined as above. If $m_{i,n}(x)$ is defined by (2.9) with $h \to 0$, $nh \to \infty$,

then

$$\int_{x_i}^{\alpha_i} \frac{xp_i'(x)}{p_i^2(x)} m_{i,n}(x)dx - \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)} dm_{i,n}(x)$$
$$\to \int_{x}^{\alpha_i} xc_i(x)dG_i(x) \text{ in } (p).$$

Proof: See Appendix B.

Now, let

$$\Delta_{i,n}^*(x_i) = \frac{\theta_0 m_{i,n}(x_i)}{p_i(x_i)} + \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)} dm_{i,n}(x) - \int_{x_i}^{\alpha_i} \frac{xp_i'(x)}{p_i^2(x)} m_{i,n}(x)dx, \quad (4.1)$$

then $\delta_n^*(\underset{\sim}{x}) = \{i \mid x_i \geq \theta_0\} \cup \{i \mid x_i < \theta_0 \text{ and } \Delta_{i,n}^*(x_i) \leq 0\}$ (4.2)

defines an empirical Bayes rule.

The following lemma is a direct result of Lemma 3 of Van Ryzin and Susarla [1977].

Lemma 4.2. Let $\Delta_{G_i}(x) = \int_{x}^{\alpha_i} (\theta_0 - t)c_i(t)dG_i(t)\, I_{(0,\alpha_i)}(x)$,

then $0 \leq r_n(\underset{\sim}{G},\underset{\sim}{\delta}_n^*) - r(\underset{\sim}{G}) = \sum_{i=1}^{k} \{\int_{H_i^1} |\Delta_{G_i}(x)| p_i(x) | P[\Delta_{i,n}^*(x) < 0]dx$

$$+ \int_{H_i^2} |\Delta_{G_i}(x)| p_i(x) P[\Delta_{i,n}^*(x) \geq 0]dx\}$$

where $\Delta^*_{i,n}(x)$ and δ^*_n are defined by (4.1) and (4.2) respectively, and $H^1_i = \{x|x < \theta_0$ and $\Delta_{G_i}(x) > 0\}$ and $H^2_i = \{x|x \le \theta_0$ and $\Delta_{G_i}(x) < 0\}$.

Now, let $O(\alpha_n)$ denote a quantity such that $0 \le \lim_{n\to\infty} \frac{O(\alpha_n)}{\alpha_n} < \infty$. Then since $|\Delta_{G_i}(x)|p_i(x) \le M_i$ for all $x \le \theta_0$ for some constant M_i, so

$$r_n(G,\delta^*_n) - r(G) \le \sum_{i=1}^{k} M_i\{\int_{H^1_i} p[\Delta^*_{i,n}(x) < 0]dx$$

$$+ \int_{H^2_i} P[\Delta^*_{i,n}(x) \ge 0]dx\}.$$

Therefore, if for all $x \le \theta_0$

$$P[|\Delta^*_{i,n}(x)-\Delta_{G_i}(x)| > |\Delta_{G_i}(x)|] = O(\alpha_n) \text{ as } n \to \infty$$

then

$$r_n(G,\delta^*_n) - r(G) = O(\alpha_n).$$

Now, by the inequality

$$P[|\Delta^*_{i,n}(x)-\Delta_{G_i}(x)| > |\Delta_{G_i}(x)|] \le \frac{\text{Var}[\Delta^*_{i,n}(x)]}{[|\Delta_{G_i}(x)|-|\Delta_{G_i}(x)-E\Delta^*_{i,n}(x)|]^2},$$

we conclude that if $\text{Var}[\Delta^*_{i,n}(x)] = O(\alpha_n)$ for all $x \le \theta_0$ then $r_n(G,\delta^*_n) - r(G) = O(\alpha_n)$.

In the following, we have carried out some Monte Carlo studies to see how fast the derived empirical Bayes rules converge. We let $X_i \sim U(0,\theta_i)$ for $i = 0,1$. θ_0 is treated as an unknown. Assume that $g_i(\theta) = \frac{2\theta}{c^2} I_{(0,c)}(\theta)$ for $i = 0,1$ and $L_1 = L_2 = 1$. The smallest sample size N such that

$$\text{Relative error} = \frac{|r_m(G,\delta^*_m)-r(G)|}{r(G)} \le \varepsilon$$

for N-4 $\leq$ m $\leq$ N is determined. The values of N corresponding to selected ε and c are shown in the next table for $h = n^{-1/4}$, for $h = n^{-1/5}$ and for $h = n^{-1/6}$, where h is used to define (2.9).

Table

Lists of values of the smallest N such that $\frac{|r_m(G,\delta^*_m)-r(G)|}{r(G)} \leq \varepsilon$ for $N-4 \leq m \leq N$, where the density of the priors is $g_i(\theta) = \frac{2\theta}{c^2} I_{(0,c)}(\theta)$ for $i = 0,1$.

$h = n^{-1/4}$						
c \ ε	.25	.20	.15	.10	.05	.01
1/3	9	10	15	25	41	-
1/2	11	12	13	14	29	-
1	15	21	25	27	86	-
2	45	60	80	122	187	-
3	61	172	174	360	-	-

$h = n^{-1/5}$						
c \ ε	.25	.20	.15	.10	.05	.01
1/3	11	13	15	21	27	-
1/2	10	13	15	21	48	-
1	13	19	20	21	46	-
2	26	27	52	151	262	-
3	51	88	134	232	304	-

$h = n^{-1/6}$						
c \ ε	.25	.20	.15	.10	.05	.01
1/3	9	10	15	25	41	-
1/2	11	12	13	14	29	-
0	11	15	20	27	97	-
1	19	31	59	60	212	-
2	51	61	136	171	302	-

Note: "-" means that N > 400 (Monte Carlo study was curtailed because of limited resources).

Appendix A

Proof of Theorem 3.1.

For i fixed, $\int_0^\infty x m_{i,n}(x) dm_{0,n}(x)$

$$= \frac{1}{n^2}\frac{1}{h^2}\sum_{j=1}^{n}\sum_{\ell=1}^{n}\int_a^\infty x\, I_{(x,x+h]}(Y_{ij})\,dI_{[Y_{0\ell}-h,Y_{0\ell})}(x)$$

$$= \frac{1}{n^2}\frac{1}{h^2}\sum_{j=1}^{n}\sum_{\ell=1}^{n}(U_{j\ell}-V_{j\ell}), \text{ where}$$

$$U_{j\ell} = (Y_{0\ell}-h)I_{(a,\infty)}(Y_{0\ell}-h)I_{(Y_{0\ell}-h,Y_{0\ell}]}(Y_{ij})$$

$$V_{j\ell} = Y_{0\ell}I_{(a,\infty)}(Y_{0\ell})I_{(Y_{0\ell},Y_{0\ell}+h]}(Y_{ij}).$$

Since $Y_{0\ell} \sim M_0(x)$ and $Y_{ij} \sim M_i(x)$ for $1 \le j, \ell \le n$, so

$$E\int_a^\infty x m_{i,n}(x) dm_{0,n}(x) = \frac{1}{h^2}E[U_{11}-V_{11}]$$

$$= \int_a^\infty x \frac{1}{h}\int_x^{x+h} dM_i(y)\,\frac{1}{h}[m_0(x+h)-m_0(x)]dx.$$

Now, by (2.2) $m_i(x)$ is decreasing in x, hence

$$\frac{1}{h}\int_x^{x+h} dM_i(y) \le m_i(x) \le \frac{1}{x}[1-G_i(x)]. \tag{A.1}$$

Then $$\left|x\cdot \frac{1}{h}\int_x^{x+h} dM_i(y)\,\frac{1}{h}[m_0(x+h)-m_0(x)]\right|$$

$$\le [1-G_i(x)]\frac{1}{h}\int_x^{x+h}\frac{1}{\theta}\,dG_0(\theta) \le \frac{1}{x}g_0(x+\delta h), \text{ for some } \delta \in [0,1].$$

The last term is integrable over (a,∞), then by LDCT

$$E\int_a^\infty xm_{i,n}(x)dm_{0,n}(x) \to \int_a^\infty xm_i(x)m_0'(x)dx$$

$$= -\int_a^\infty m_i(x)dG_0(x) \text{ in (p) if } h \to 0 \text{ as } n \to \infty. \tag{A.2}$$

Now, $\operatorname{Var}\int_a^\infty xm_{i,n}(x)dm_{0,n}(x) = \operatorname{Var}\frac{1}{n^2}\frac{1}{h^2}\sum_{j,\ell}(U_{j\ell}-V_{j\ell})$

$$= \frac{1}{n^2h^4}\operatorname{Var}(U_{11}-V_{11}) + \frac{2(n-1)}{n^2h^4}\operatorname{Cov}(U_{11}-V_{11},\ U_{12}-V_{12}). \tag{A.3}$$

But $\operatorname{Var}(U_{11}-V_{11}) \le E[(U_{11}-V_{11})^2] = E(U_{11}^2) + E(V_{11}^2)$ [because $U_{11}V_{11} = 0$], and $\frac{1}{h}E(U_{11}^2)$

$$= \int_a^\infty x^2\cdot\frac{1}{h}\int_x^{x+h} dM_i(y)dM_0(x+h)$$

$$< \int_a^\infty x^2\cdot\frac{1}{x}(1-G_i(x))dM_0(x+h) \le \int_a^\infty x\ dM_0(x+h)$$

$$\le E^{M_0}[X] = E^{G_0}[E[X|\theta_0]] = \frac{1}{2}E^{G_0}[\theta_0] < \infty,$$

hence $\frac{1}{h}\operatorname{Var}(U_{11}-V_{11}) \le E^{G_0}[\theta_0]$ for all $h > 0$. (A.4)

Meanwhile, $\operatorname{Cov}(U_{11}-V_{11},\ U_{12}-V_{12}) = \operatorname{Cov}(U_{11},U_{12}) + \operatorname{Cov}(V_{11},V_{12}) - \operatorname{Cov}(U_{11},V_{12}) - \operatorname{Cov}(V_{11},U_{12})$, and $|\frac{1}{h^2}\operatorname{Cov}(U_{11},U_{12})| \le \frac{1}{h^2}[E(U_{11}U_{12}) + E(U_{11})E(U_{12})]$ because $U_{j\ell} \ge 0$ for all $1 \le j,\ \ell \le n$.

Now, $\frac{1}{h^2}E(U_{11}U_{12}) = \frac{1}{h^2}\int_0^\infty [\int_{(a,\infty)\cap[x-h,x)} ydM_0(y+h)]^2 dM_i(x)$

$$= \frac{1}{h^2}\int_{a+h}^\infty [\int_{x-h}^x ydM_0(y+h)]^2 dM_i(x) + \frac{1}{h^2}\int_a^{a+h}[\int_a^x ydM_0(y+h)]^2 dM_i(x).$$

Because $\int_{x-h}^x ydM_0(y+h) = \int_{x-h}^x y\int_{y+h}^\infty \frac{1}{\theta}dG_0(\theta)dy$

$$\le \int_{x-h}^x y\cdot\frac{1}{y+h}\cdot dy \le h, \text{ and, similarly,}$$

$$\int_a^x y dM_0(y+h) < \int_a^{a+h} y dM_0(y+h) \le h \quad \text{for} \quad a < x < a+h,$$

we get $\frac{1}{h^2} E(U_{11}U_{12}) \le 1-M_i(a+h) + M_i(a+h) - M_i(a) = 1-M_i(a)$.
The same argument shows that $\frac{1}{h} E(U_{11}) \le 1-M_i(a)$

$$\frac{1}{h} E(V_{11}) \le 1-M_i(a),$$

hence $|\frac{1}{h^2} \mathrm{Cov}(U_{11},U_{12})| \le 2[1-M_i(a)]$. This implies that

$$\frac{1}{h^2} |\mathrm{Cov}(U_{11}-V_{11},U_{12}-V_{12})| \le 8\,[1-M_i(a)] \text{ for any } h > 0. \tag{A.5}$$

By (A.3), (A.4) and (A.5)

$$\mathrm{Var} \int_a x m_{i,n}(x) dm_{0,n}(x) \to 0 \quad \text{if} \quad nh^2 \to 0 \text{ and } h \to 0. \tag{A.6}$$

Now, (A.2) and (A.6) implies that

$$\int_a^\infty x m_{i,n}(x) dm_{0,n}(x) \to - \int_a^\infty m_i(x) dG_0(x) \text{ in } (p).$$

This finishes the proof.

Appendix B

Proof of Theorem 4.1.

$$\text{First,}\quad E\int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_{i,n}(x) = \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)} \frac{1}{h}\,[m_i(x+h)-m_i(x)]dx$$

$$\to \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_i(x) \quad \text{by LDCT.}$$

$$\text{Now,}\quad \mathrm{Var}\int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_{i,n}(x) = \mathrm{Var}\Big[\frac{1}{nh}\sum_{j=1}^{n}(U_j-V_j)\Big]$$

$$\text{where}\quad U_j = \frac{Y_{ij}-h}{p_i(Y_{ij}-h)}\, I_{[x_i,\alpha_i]}(Y_{ij}-h), \text{ and}$$

$$V_j = \frac{Y_{ij}}{p_i(Y_{ij})}\, I_{[x_i,\alpha_i]}(Y_{ij}).$$

$$\text{Hence,}\quad \mathrm{Var}\int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_{i,n}(x) = \frac{1}{nh^2}\,\mathrm{Var}(U_1-V_1)$$

$$\leq \frac{1}{nh^2}\, E[(U_1-V_1)^2] = \frac{1}{n}\int_{x_i+h}^{\alpha_i} \Big[\frac{1}{h}\Big(\frac{x}{p_i(x)} - \frac{x-h}{p_i(x-h)}\Big)\Big]^2 dM_i(x)$$

$$+ \frac{1}{nh}\int_{\alpha_i}^{\alpha_i+h} \frac{1}{h}\Big[\frac{x-h}{p_i(x-h)}\Big]^2 dM_i(x) + \frac{1}{nh}\int_{x_i}^{x_i+h} \frac{1}{h}\,\frac{x^2}{p_i^2(x)}\, dM_i(x)$$

$$\leq \frac{1}{n}\max_{x\in[x_i,\alpha_i]} \Big[\frac{d}{dx}\,\frac{x}{p_i(x)}\Big]^2 + \frac{2}{nh}\max_{x\in[x_i,\alpha_i]} \Big[\frac{x}{p_i(x)}\Big]^2$$

$$\to 0 \quad \text{if} \quad nh \to \infty.$$

We see that

$$\int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_{i,n}(x) \to \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)}\, dm_i(x) \quad \text{in (p).}$$

Similarly $\int_{x_i}^{\alpha_i} \frac{xp_i'(x)}{p_i^2(x)} m_{i,n}(x)dx \to \int_{x_i}^{\alpha_i} \frac{xp_i'(x)}{p_i^2(x)} m_i(x)dx$ in (p).

Since $\int_{x_i}^{\alpha_i} xc_i(x)dG_i(x) = \int_{x_i}^{\alpha_i} -x \frac{d}{dx} [\frac{m_i(x)}{p_i(x)}]$

$$= \int_{x_i}^{\alpha_i} \frac{xp_i'(x)}{p_i^2(x)} m_i(x)dx - \int_{x_i}^{\alpha_i} \frac{x}{p_i(x)} dm_i(x),$$

the proof is completed.

Bibliography

Barr, D. R., and Rizvi, M. H. (1966). Ranking and selection problems of uniform distributions. Trabajos Estadist., 17, 15-31.

Deely, J. J. (1965). Multiple Decision Procedures from an Empirical Bayes Approach. Ph.D. Thesis (Mimeo. Ser. No. 45), Dept. of Statist., Purdue Univ., West Lafayette, Indiana 47907.

Fox, R. J. (1978). Solution to empirical Bayes squared error loss estimation problems. Ann. Statist., 6, 846-854.

Gupta, S. S., and Sobel, M. (1958). On selecting a subset which contains all populations better than a control. Ann. Math. Statist., 29, 235-244.

Gupta, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math., 14, 199-216.

Gupta, S. S. (1965). On some multiple decision (selecting and ranking) rules. Technometrics, 7, 225-245.

Van Houwelingen, J. C. (1976). Montone empirical Bayes tests. Ann. Statist., 4, 981-989.

Hwang, W. T. (1975). Bayes approach to a problem of partitioning k normal populations. Bull. Inst. Math. Acad. Sinica, 3, 87-97.

McDonald, G. C. (1974). The distribution of a variate based on independent ranges from a uniform population. Tech. Report GMR-1775, General Motors Research Laboratories, Warren, Michigan.

Meeden, G. (1972). Some admissible empirical Bayes procedures. Ann. Math. Statist., 43, 96-101.

Parzen, E. (1962). On estimation of a probability density function and model. Ann. Math. Statist., 33, 1065-1076.

Robbins, H. (1955). An empirical Bayes approach to statistics. Proc. 3rd Berkeley Symp. Math. Prob. University of California Press, 155-163.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. Ann. Math. Statist., 35, 1-19.

Van Ryzin, J. (1970). Empirical Bayes procedures for multiple decision problems. Tech. Report No. 249, Dept. of Statist., Univ. of Wisconsin, Madison.

Van Ryzin, J., and Susarla, V. (1977). On the empirical Bayes approach to multiple decision problems. Ann. Statist., 5, 172-181.

Singh, A. K. (1977). On slippage test and multiple decision (selection and ranking) procedures. Ph.D. Thesis. Dept. of Statist., Purdue Univ., West Lafayette, Indiana 47907.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Mimeograph Series #81-5
2. GOVT ACCESSION NO.: AD-A097 619
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Empirical Bayes Rules for Selecting Good Populations
5. TYPE OF REPORT & PERIOD COVERED: Technical Rept.
6. PERFORMING ORG. REPORT NUMBER: Mimeo. Series #81-5
7. AUTHOR(s): Shanti S. Gupta ~~and~~ Ping Hsiao
8. CONTRACT OR GRANT NUMBER(s): ONR N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, West Lafayette, IN 47907
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, DC
12. REPORT DATE: Mar[redacted]81
13. NUMBER OF PAGES: 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): MIMEOGRAPH-SER-81-5
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release, distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Asymptotically optimal, empirical Bayes, truncation parameter, rate of convergence, Monte Carlo study.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A problem of selecting populations better than a control is considered. When the populations are uniformly distributed, empirical Bayes rules are derived for a linear loss function for both the known control parameter and the unknown control parameter cases. When the priors are assumed to have bounded supports, empirical Bayes rules for selecting good populations are derived for distributions with truncation parameters (i.e. the form of the pdf is $f(x|\theta) = p_i(x)c_i(\theta)I_{(0,\theta)}(x)$). Monte Carlo studies are carried out which determine the

DD FORM 1 JAN 73 1473

UNCLASSIFIED 291730

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

minimum sample sizes needed to make the relative errors less than , for given -values.

END

DATE
FILMED
5 81

DTIC